

Volume 4 Nomor 1 Januari 2019

INFORMASI INTERAKTIF

JURNAL INFORMATIKA DAN TEKNOLOGI INFORMASI

PROGRAM STUDI TEKNIK INFORMATIKA – FAKULTAS TEKNIK -UNIVERSITAS JANABADRA

PENERAPAN ALGORITMA *K NEAREST NEIGHBOR* UNTUK REKOMENDASI MINAT KONSENTRASI DI PROGRAM STUDI TEKNIK INFORMATIKA UNIVERSITAS PGRI YOGYAKARTA

Adi Prasetyo, Kusri, M. Rudyanto Arief

DETEKSI GEJALA VIRUS ZIKA MENGGUNAKAN *CERTAINTY FACTOR* DAN *NAIVE BAYES* BERBASIS ANDROID

Emma Nur Hamidah, Ryan Ari Setyawan, Fatsyahrina Fitriastuti

KLASIFIKASI JENIS REMPAH-REMPAH BERDASARKAN FITUR WARNA RGB DAN TEKSTUR MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

Kaharuddin, Kusri, Emha Taufiq Luthfi

PENERAPAN ALGORITMA PALGUNADI PADA *SPLIT DELIVERY VEHICLE ROUTING PROBLEM* UNTUK PENDISTRIBUSIAN MULTI PRODUK

Sri Wulandari, Kusri, M. Rudyanto Arief

PEMODELAN ARSITEKTUR SISTEM INFORMASI PERHOTELAN DENGAN KERANGKA KERJA TOGAF ADM

Selviana Yunita, Wing Wahyu Winarno, Asro Nasriri

PERANCANGAN SISTEM PEMANTAUAN WAKTU NYATA BERBASIS *INTERNET OF THINGS (IOT)*

Mat Sudir, Bambang Soedjono W A, Eko Pramono

PENERAPAN DATA MINING DALAM MENENTUKAN PEMBINAAN KOPERASI (STUDI KASUS : DINAS KOPERASI DAN UKM KABUPATEN KOTAWARINGIN TIMUR)

Yuni Ambar S, Kusri, Henderi

IMPLEMENTASI *DATABASE SECURITY* MENGGUNAKAN KONSEP *ROLE-BASED ACCESS CONTROL (RBAC)* DALAM RANCANGAN DATABASE SISTEM INFORMASI MANAJEMEN SEKOLAH DENGAN POSTGRESQL

Achmad Yusron Arif, Emma Utami, Suwanto Raharjo

RANCANG BANGUN VISUALISASI *TOURISM GUIDE* PROVINSI DAERAH ISTIMEWA YOGYAKARTA

Jeffry Andhika Putra, Rusdy Agustaf



INFORMASI
INTERAKTIF

Vol. 4

No. 1

Hal. 1 - 62

Yogyakarta
Januari 2019

ISSN
2527-5240

DEWAN EDITORIAL

- Penerbit** : Program Studi Teknik Informatika Fakultas Teknik Universitas Janabadra
- Ketua Penyunting
(Editor in Chief)** : Fatsyahrina Fitriastuti, S.Si., M.T. (Universitas Janabadra)
- Penyunting (Editor)** : 1. Selo, S.T., M.T., M.Sc., Ph.D. (Universitas Gajah Mada)
2. Dr. Kusriani, S.Kom., M.Kom. (Universitas Amikom Yogyakarta)
3. Jemmy Edwin B, S.Kom., M.Eng. (Universitas Janabadra)
4. Ryan Ari Setyawan, S.Kom., M.Eng. (Universitas Janabadra)
5. Yumarlin MZ, S.Kom., M.Pd., M.Kom. (Universitas Janabadra)
- Alamat Redaksi** : Program Studi Teknik Informatika Fakultas Teknik
Universitas Janabadra
Jl. Tentara Rakyat Mataram No. 55-57
Yogyakarta 55231
Telp./Fax : (0274) 543676
E-mail: informasi.interaktif@janabadra.ac.id
Website : <http://e-journal.janabadra.ac.id/>
- Frekuensi Terbit** : 3 kali setahun

JURNAL INFORMASI INTERAKTIF merupakan media komunikasi hasil penelitian, studi kasus, dan ulasan ilmiah bagi ilmuwan dan praktisi dibidang Teknik Informatika. Diterbitkan oleh Program Studi Teknik Informatika Fakultas Teknik Universitas Janabadra di Yogyakarta, tiga kali setahun pada bulan Januari, Mei dan September.

DAFTAR ISI

	<i>halaman</i>
Penerapan Algoritma <i>K Nearest Neighbor</i> Untuk Rekomendasi Minat Konsentrasi Di Program Studi Teknik Informatika Universitas PGRI Yogyakarta Adi Prasetyo, Kusriani, M. Rudyanto Arief	1 – 6
Deteksi Gejala Virus Zika Menggunakan <i>Certainty Factor</i> dan <i>Naive Bayes</i> Berbasis Android Emha Nur Hamidah, Ryan Ari Setyawan, Fatsyahrina Fitriastuti	7 – 16
Klasifikasi Jenis Rempah-Rempah Berdasarkan Fitur Warna RGB dan Tekstur Menggunakan Algoritma <i>K-Nearest Neighbor</i> Kaharuddin, Kusriani, Emha Taufiq Luthfi	17 – 22
Penerapan Algoritma Palgunadi pada <i>Split Delivery Vehicle Routing Problem</i> untuk Pendistribusian Multi Produk Sri Wulandari, Kusriani, M.Rudyanto Arief	23 – 30
Pemodelan Arsitektur Sistem Informasi Perhotelan dengan Kerangka Kerja TOGAF ADM Selviana Yunita, Wing Wahyu Winarno, Asro Nasriri	31 – 38
Perancangan Sistem Pemantauan Waktu Nyata Berbasis <i>Internet of Things</i> (IoT) Mat Sudir, Bambang Soedjono W A, Eko Pramono	39 – 43
Penerapan <i>Data Mining</i> dalam Menentukan Pembinaan Koperasi (Studi Kasus : Dinas Koperasi Dan UKM Kabupaten Kotawaringin Timur) Yuni Ambar S, Kusriani, Henderi	44 – 50
Implementasi <i>Database Security</i> Menggunakan Konsep <i>Role-Based Access Control</i> (RBAC) dalam Rancangan Database Sistem Informasi Manajemen Sekolah Dengan PostgreSQL Achmad Yusron Arif, Emha Utami, Suwanto Raharjo	51 – 55
Rancang Bangun Visualisasi Tourism Guide Provinsi Daerah Istimewa Yogyakarta Jeffry Andhika Putra, Rusdy Agustaf	56 - 62

PENGANTAR REDAKSI

Puji syukur kami panjatkan kehadiran Allah Tuhan Yang Maha Kuasa atas terbitnya JURNAL INFORMASI INTERAKTIF Volume 4, Nomor 1, Edisi Januari 2019. Pada edisi kali ini memuat 9 (sembilan) tulisan hasil penelitian dalam bidang teknik informatika.

Harapan kami semoga naskah yang tersaji dalam JURNAL INFORMASI INTERAKTIF edisi Januari tahun 2019 dapat menambah pengetahuan dan wawasan di bidangnya masing-masing dan bagi penulis, jurnal ini diharapkan menjadi salah satu wadah untuk berbagi hasil-hasil penelitian yang telah dilakukan kepada seluruh akademisi maupun masyarakat pada umumnya.

Redaksi

PENERAPAN ALGORITMA *K NEAREST NEIGHBOR* UNTUK REKOMENDASI MINAT KONSENTRASI DI PROGRAM STUDI TEKNIK INFORMATIKA UNIVERSITAS PGRI YOGYAKARTA

*Adi Prasetyo*¹, *Kusrini*², *M Rudyanto Arief*³

^{1,2,3}Universitas AMIKOM Yogyakarta

Jl. Ringroad Utara, Condongcatur, Depok, Sleman, Yogyakarta Indonesia 55283

Email : ¹⁾*adipras@upy.ac.id*, ²⁾*kusrini@amikom.ac.id*³⁾*rudy@amikom.ac.id*

ABSTRACT

The graduation time of a student is very important because it relates to many parties, in addition to the students concerned, guardian lecturers, and program study programs and related parties in the time of student graduation. one of the factors which is the result of graduation from informatics engineering study program students is that many students are wrong in taking thesis titles that are not in accordance with their interests and concentration. So research needs to be done to solve this problem. One way to solve problems is to make predictions. Selection of interest by using student data using Case Base Reasoning (CBR). In this study applying the K-NN Algorithm in determining recommendations in choosing the concentration. The results obtained from this study are by classification using the K-NN Algorithm obtained concentrations that match their interests.

Keywords : *Predictions, Case Base Reasoning, K-NN Algorithm*

1. PENDAHULUAN

Waktu kelulusan seorang mahasiswa merupakan hal yang sangat penting karena berhubungan dengan banyak pihak, selain mahasiswa yang bersangkutan, dosen wali, dan ketua program studi serta pihak-pihak yang terkait dalam waktu kelulusan mahasiswa. Lama waktu kelulusan mahasiswa dapat digunakan untuk evaluasi penyelenggaraan pendidikan. Pada Akreditasi Program Studi Buku V Pedoman Penilaian Instrumen Akreditasi (BAN-PT, 2009) standar 3, yaitu tentang mahasiswa dan kelulusan. Salah satu elemen penilaiannya adalah rata-rata masa studi lulusan (dalam tahun). Hal ini membuktikan bahwa kelulusan seorang mahasiswa merupakan hal penting dalam akreditasi.

Dilakukannya prediksi pemilihan konsentrasi mahasiswa pada program studi teknik informatika UPY agar mahasiswa dapat lulus sesuai dengan target yang ditetapkan oleh Program Studi yaitu 100 lulusan mahasiswa per semester. Untuk saat ini rekomendasi minat konsentrasi sangat perlu dilakukan di lingkungan prodi teknik informatika UPY karena sampai saat ini tingkat kelulusan mahasiswa dalam satu angkatan masih belum

sesuai dengan target yang ditetapkan yaitu 100 mahasiswa per semester. Berdasarkan informasi dari hasil rapat program studi tahun akademik 2016/2017 dapat diketahui bahwa salah satu faktor yang menyebabkan masih kurangnya tingkat kelulusan mahasiswa prodi teknik informatika adalah banyak mahasiswa yang salah dalam pengambilan judul skripsi yang tidak sesuai dengan minat dan konsentrasinya. Sehingga dalam pembuatan skripsi tidak berjalan sesuai dengan yang diinginkan. Faktor tersebut dipengaruhi dari kurang tepatnya mahasiswa dalam memilih konsentrasi minat yang ada. Tidak tepatnya dalam pemilihan konsentrasi akan berdampak langsung bagi mahasiswa. Berdasarkan dari permasalahan dari kasus yang sudah dijelaskan diatas, maka perlu dilakukan penelitian untuk memecahkan masalah tersebut. Salah satu cara untuk memecahkan permasalahan tersebut adalah melakukan prediksi pemilihan minat konsentrasi dengan cara menggali data pengalaman para mahasiswa menggunakan *Case Base Reasoning* (CBR). Sedangkan algoritma yang digunakan adalah algoritma *Euclidean dan hamming distance* karena akan dilakukan proses pencocokan terhadap kasus-kasus lama yang telah tersimpan dalam basis

kasus. Pengukuran kemiripan kasus baru dengan kasus lama dilakukan menggunakan algoritma *nearest neighbor* yang kemudian dicari tingkat kemiripan lokalnya (*similarity local*) antara basis kasus dengan kasus yang akan diprediksi setiap fitur dengan menggunakan *euclidean distance* untuk fitur inputan berupa numerik dan *hamming distance* untuk fitur inputan berupa huruf. Metode perhitungan *similarity* ini menghitung jarak terdekat antara fakta pada kasus yang akan diprediksi dengan kasus yang sudah ada. Setelah setiap fitur dihitung tingkat kemiripan lokalnya, kemudian dilakukan perhitungan tingkat kemiripan secara keseluruhan (*similarity global*). Masing-masing fitur memiliki bobot kepentingan sendiri-sendiri. Hasil dari *similarity global* ini yang menentukan seberapa dekat antara basis kasus dengan kasus yang akan diprediksi.

2. TINJAUAN PUSTAKA

Penelitian yang dilakukan oleh Khairul Hafidh (2015), melakukan prediksi masa studi mahasiswa menggunakan *Case-Based Reasoning* dengan metode *Jaccard Coefficient* dengan hasil akurasi dari 31 kasus sebesar 80.65% [1]. CBR digunakan untuk menyelesaikan kasus baru dengan mengingat situasi yang pernah terjadi dengan mengambil solusi yang baru untuk kasus yang mirip. CBR tidak memerlukan suatu model yang eksplisit dan pengetahuan didapatkan dengan cara mengumpulkan kejadian-kejadian yang telah terjadi. Representasi kasusnya hanya menggunakan 2 komponen yaitu problem dan solusi. *Indexing* juga diterapkan untuk mempercepat dalam penelusuran data pada saat pencarian.

Penelitian lain yang dilakukan oleh Rohmadi (2015) dengan judul *Case Based Reasoning* untuk pemilihan kegiatan organisasi mahasiswa, penelitian ini bertujuan untuk memberikan saran atau masukan bagi mahasiswa baru dalam mengambil kegaitana organisasi mana yang cocok dengan mereka [2]. Wahyudi (2015) juga melakukan penelitian menggunakan CBR untuk mendiagnosis penyakit jantung [3]. Pada proses diagnosis dilakukan dengan cara memasukkan gejala baru ke dalam sistem, kemudian melakukan proses perhitungan nilai similaritas antara permasalahan baru dengan basis kasus menggunakan metode *nearest neighbor similarity*, *minskowski distance similarity*, dan

euclidean distance similarity. Jika suatu kasus tidak berhasil diagnosis atau target case <80, maka akan dilakukan revisi kasus oleh pakar. Hasil pengujian rata-rata sistem untuk melakukan diagnosis awal terhadap penyakit jantung menggunakan data rekam medik pasien jantung, menunjukkan bahwa sistem mampu mengenali penyakit jantung secara benar sebesar 100% dengan menggunakan ketiga metode tersebut. Serta perhitungan tingkat akurasi menggunakan metode *nearest neighbor similarity* sebesar 86,21%, metode *minskowski* sebesar 100% dan metode *euclidean* sebesar 94,83%.

Berikutnya adalah penelitian yang dilakukan oleh Wita Yulianti (2016) tentang Penerapan *Case Based Reasoning* dalam sistem pakar untuk mementukan minat dan bakat pada siswa Sekolah Dasar (SD), penelitian ini menggunakan algoritma *K Nearest Neighbour* yang bertujuan untuk menentukan minat dan bakat siswa sekolah dasar, lebih diutamakan untuk siswa kelas IV, V, VI [4]. Dikarenakan siswa kelas tersebut telah memasuki usia sekitar 10-11 tahun, dimana saat usia tersebut adalah usia peralihan ke dunia nyata ketika kemampuan kognitif anak berkembang untuk merencanakan sesuatu.

Penelitian selanjutnya adalah penelitian yang akan dilakukan oleh Ulfi (2016) dengan judul *Prediksi Waktu Kelulusan Mahasiswa dengan CBR pada Pascasarjana Program Studi Ilmu Komputer di Universitas Gadjah Mada* [5]. Pada penelitian ini akan dibuat sebuah sistem yang dapat melakukan prediksi kelulusan mahasiswa S2 Ilmu Komputer menggunakan *case-based reasoning* yang diharapkan dapat lebih mudah beradaptasi dengan kasus-kasus baru dengan melakukan pencocokan terhadap kasus-kasus lama yang telah tersimpan dalam basis kasus. Pencocokan yang dilakukan pada penelitian ini menggunakan *Euclidean distance* dan *Hamming distance*. Penelitian ini diharapkan dapat meningkatkan akurasi dalam melakukan prediksi waktu kelulusan mahasiswa S2 Ilmu Komputer dibandingkan dengan menggunakan *Jaccard Coefficient* seperti penelitian yang dilakukan oleh Hafidh (2015).

3. LANDASAN TEORI

3.1 Case Based Reasoning

Case Based Reasoning (CBR) atau penalaran berbasis kasus merupakan suatu teknik pemecahan masalah yang mengadopsi

solusi masalah-masalah sebelumnya yang mirip dengan masalah baru yang dihadapi untuk mendapatkan solusinya (Riesbeck dan Shank, 1994). Kasus-kasus pada masa lalu disimpan dengan menyertakan fitur-fitur yang menggambarkan karakteristik dari kasus tersebut beserta solusinya.

CBR telah banyak digunakan di berbagai bidang. Secara umum aplikasi CBR dikelompokkan menjadi kelompok klasifikasi dan kelompok sintesis. Pada kelompok klasifikasi, setiap kasus dikelompokkan berdasarkan kemiripannya yang dilihat dari fitur-fiturnya. Kasus yang baru akan diadopsi solusi dari kasus-kasus yang berada dalam kelompok kasus yang sama. Pada kelompok sintesis, solusi didapat dengan cara mengkombinasikan solusi-solusi dari kasus-kasus sebelumnya. Kelompok ini biasanya merupakan suatu sistem gabungan antara CBR dengan teknik yang lain.

3.2 Tipe Tipe Aplikasi *Case Based Reasoning*

Aplikasi CBR terbagi menjadi dua tipe permasalahan, yaitu :

a. Tipe Klasifikasi

Sebuah kasus baru akan dicocokkan dalam basis kasus untuk menentukan tipe, atau kelas dari kasus tersebut. Solusi dari pencocokan kasus terbaik dalam kelas kemudian digunakan kembali.

b. Kemudahan Tipe Klasifikasi

Tipe klasifikasi mudah diimplementasikan, karena :

1. Tipe klasifikasi sangat sesuai dengan siklus /tahap sistem penalaran komputer berbasis kasus
2. Kasus cenderung lebih mudah direpresentasikan dan mudah dikumpulkan. Algoritma penelusuran digunakan kebanyakan tool sistem penalaran komputer berbasis kasus adalah klasifikasi.

c. Jenis Tipe Klasifikasi

Tipe klasifikasi terdiri dari beberapa jenis, yaitu :

1. Diagnosa, contoh : diagnosa penyakit.
2. Prediksi , contoh : peramalan saham.
3. Penaksiran, contoh : analisis resiko untuk bank

4. Kendali proses, contoh : rencan perjalanan, atau penjadwalan kerja.

d. Tipe Sintesis

Mencoba untuk membuat solusi baru dengan menggabungkan bagian dari solusi sebelumnya. Tipe sintesis sudah menjadi sifat kompleks karena batasan antara elemen-elemen yang digunakan selama sintesis. Sistem penalaran menggunakan adaptasi dan biasanya sistem hybrid menggabungkan sistem penalaran komputer berbasis kasus dengan teknik lain.

e. Jenis Tipe Sintesis

Ada beberapa jenis tipe sintesis, yaitu

1. Desain, membuat bagian baru dengan adaptasi elemen dari yang sebelumnya.
2. Perencanaan, membuat perencanaan baru dari elemen perencanaan sebelumnya
3. Konfigurasi, membuat penjadwalan baru dari penjadwalan sebelumnya.

f. Alasan Jenis Tipe Sintesis Sulit Dibangun

Beberapa alasan mengapa sistem penalaran komputer berbasis kasus sintesis sulit dibangun :

1. Representasi kasus dari sebuah perencanaan atau desain seringkali kompleks dan menggunakan struktur ringgi dengan banyak fitur yang saling berkaitan.
2. Kasus-kasus biasanya tidak disimpan dalam media tunggal, pengumpulan kasus menjadi lebih sulit.
3. Adaptasi sering kebutuhan utama dalam tipe sintesis.
- 4.

g. Teknik *Case Based Reasoning*

Teknik-teknik yang digunakan dalam CBR yaitu : representasi kasus, *indexing*, *storage* (penyimpanan), *retrieval*, dan adaptasi

h. Representasi Kasus

Penalaran berbasis kasus bergantung pada struktur da nisi dari koleksi kasus. Suatu kasus dapat di selesaikan dengan memanggil kasus sebelumnya yang sesuai / cocok dengan kasus yang baru. Sebuah kasus dapat menjadi sebuah catatan dari kejadian atau record secara khusus terdiri dari :

1. Permasalahan (*problem*) yang menjelaskan keadaan nyata ketika kasus terjadi
2. Solusi keadaan diperoleh seolusi permasalahan

Pada gambar 1 dapat dilihat bahwa kasus tunggal yang terdiri dari dua komponen yaitu : deskripsi permasalahan (*problem description*) dan solusi telah disimpan (*solution stored*). Deskripsi permasalahan terletak di *problem space*, dan solusi tersimpan di *solution space*. Deskripsi dari kasus yang baru akan diselesaikan diletakkan di dalam *problem space*. Penelusuran mengidentifikasi kasus yang paling mirip dengan deskripsi permasalahan (panah berlabel “R”) dan solusi yang tersimpan ditemukan. Jika perlu, adaptasi dilakukan (panah berlabel “A”) dan solusi baru dibuat. Model konseptual penalaran computer berbasis kasus ini diasumsikan bahwa ada hubungan satu ke satu antara *problem* dan *solution space*

3.3 K Nearest Neighbor

K-Nearest-Neighbor (KNN) adalah salah satu metode dimana metode ini melakukan klasifikasi berdasarkan data training atau data pembelajaran dilihat dari jarak yang paling dekat dengan objek berdasarkan nilai k. Metode ini bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample. Diberikan suatu titik query, selanjutnya akan ditemukan sejumlah K objek atau titik training yang paling dekat dengan titik query. Nilai prediksi dari query akan ditentukan berdasarkan klasifikasi tetangga. Sebelum melakukan perhitungan dengan metode K-Nearest Neighbor, terlebih dahulu harus menentukan data latih dan data uji. Kemudian akan dilakukan proses perhitungan untuk mencari jarak menggunakan Euclidean. Teknik ini sangat sederhana dan mudah diimplementasikan. Mirip dengan teknik clustering, yaitu mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga terdekat. Pertama, sebelum mencari jarak data ke tetangga adalah menentukan nilai K tetangga (neighbor). Lalu, untuk mendefinisikan jarak antara dua titik yaitu titik pada data training dan titik pada data testing, maka digunakan rumus Euclidean.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (1)$$

Keterangan :

$d(x_i, x_j)$: Jarak Euclidean (Euclidean Distance).

$(x_i), (x_j)$: record ke-i , record ke-j
 (a_r) : data ke-r
 i, j : 1,2,3,...n
 n : dimensi objek

4. METODE PENGUMPULAN DATA

Dalam melakukan penelitian ini, akan digunakan cara-cara penelitian sebagai berikut :

1. Studi Pustaka

Merupakan metode pengumpulan data dengan cara mencari, membaca dan mengumpulkan dokumen dokumen sebagai referensi seperti buku, artikel dan literatur yang berhubungan dengan *Case Based Reasoning, K Nearest Neighbor, Encludian and Hamming Distance*.

2. Wawancara

Metode yang dilakukan dengan cara mengajukan pertanyaan atau Tanya jawab kepada Ketua Program Studi dan para mahasiswa angkatan 2015 sampai 2017. Metode ini digunakan supaya memastikan bahwa data yang di peroleh adalah data yang valid.

3. Observasi

Metode ini adalah metode yang menggunakan cara mempelajari dan mengetahui obyek yang akan di teliti secara langsung. Metode ini juga bisa meliputi untuk mendapatkan data dan informasi yang di butuhkan diperoleh dari data mahasiswa yang ada di Program Studi Teknik Informatika. Data yang diperoleh merupakan data sekunder dari Biro Akademik (BAAK) Universitas PGRI Yogyakarta yang mempunyai wewenang dan membunyai kebijakan dalam proses oleh data mahasiswa

5. HASIL DAN PEMBAHASAN

Sebelum melakukan proses *Case Based Reasoning* harus ditentukan dulu hasil dari perhitungan K-NN data yang digunakan dalam penelitian ini yaitu sebanyak 100 data mahasiswa tahun 2017 yang akan menentukan konsentrasi jurusan masing-masing, data tersebut akan dibagi ke dalam data training dan data testing. Pembagian dilakukan dengan persentase, yaitu 70 % pada data training dan 30% pada data testing. Setelah dilakukan proses pembagian data tersebut maka diperoleh 70 data training dan 30 data testing. Data mahasiswa berdasarkan atribut-atribut yang digunakan seperti terlihat pada Tabel 1.

Tabel 1. Data Mahasiswa

No	Nama	Jurusan SMA	UAS SMA	Nilai PMB	Berata Sem	Konsentrasi
1	Ahmad Bahag	IPS	6,3	68	2,75	Sistem Informasi
2	Yuliani Sulastri	IPS	6,4	72	2,69	Jaringan
3	Yuni Selanti	IPA	7	69	3	Sistem Cerdas
4	Prasetyo Sugiharto	IPA	6,5	75	2,5	Jaringan
5	Sigit Hartanto	IPS	6,1	82	3,2	Sistem Cerdas
6	Ansoni	IPA	6,5	77	2,3	Jaringan
7	Luluk Inani	IPS	6,2	68	3,3	Sistem Cerdas
8	Lusi Prabudi	IPA	6,5	70	3,2	Sistem Cerdas
9	Lina	IPS	7	71	2,98	Sistem Cerdas
10	Dona Rahmanto	IPA	6,7	68	2,78	Jaringan
s/d						
100	M. Ifan	IPS	6,2	68	2,66	Sistem Informasi

Setelah diperoleh data dilakukan normalisasi data yaitu dengan proses membuat skala dari nilai atribut sehingga bisa jatuh pada range tertentu. Untuk menghitung jarak menggunakan rumus *Euclidean* maka langkah berikutnya adalah mengubah data kategori ke dalam numeric. Dapat dilakukan dengan mengganti data dengan angka tertentu asalkan konsisten, Misalkan untuk atribut Jurusan untuk IPA = 1, sedangkan IPS = 2, dan setelah dilakukan normalisasi maka dapat dilihat seperti pada Tabel 2.

Tabel 2. Normalisasi

NM	JUR	UAS	PMB	SEM	Konsentrasi
A1	1	6,2	68	2,75	Sistem Informasi
A2	1	6,4	72	2,69	Jaringan
A3	0	7	69	3	Sistem Cerdas
A4	0	6,5	75	2,5	Jaringan
A5	1	6,3	82	3,2	Sistem Cerdas
A6	0	6,5	77	2,3	Jaringan
A7	1	6,2	68	3,3	Sistem Cerdas
A8	0	6,5	70	3,2	Sistem Cerdas
A9	1	7	71	2,98	Sistem Cerdas
A10	0	6,7	68	2,78	Jaringan
s/d					
A100	1	6,2	68	2,66	Sistem Informasi

Setelah didapatkan data yang telah di normalisasi kemudian menentukan nilai euclidean distance dengan nilai k yang paling akurat, untuk mencari nilai k yang paling akurat

dengan menggunakan validasi silang (cross validation). Akan dilakukan 3 kali pengujian (Q) menggunakan data training yang berbeda-beda, dan akan digunakan K=5, K=7, K=9. Pembagian data training adalah seperti pada Tabel 3.

Tabel 3. Pembagian Data Training

NO	Training (Q)	Data Testing
1	Q1	A11 s/d A40
2	Q2	A41 s/d A70
3	Q3	A71s/d A100

Setelah diperoleh data yang digunakan untuk pengujian kemudian dicari klasifikasi dari masing-masing data testing melalui pengujian Q1 s/d Q3. Untuk pengujian Q1 data testing A11 s/d A40, selanjutnya perhitungan Algoritma K-NN dengan langkah-langkah sebagai berikut :

- Menentukan parameter *k* (jumlah tetangga paling dekat).
Parameter *k* (jumlah tetangga paling dekat untuk data koperasi diatas adalah menggunakan nilai *k*=5, *k*=7, *k*=9.
- Rumus K-N dalam perhitungan kuadrat jarak euclid masing – masing objek terhadap data sample yang diberikan adalah :

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2}$$
- Kemudian hasil perhitungan pada data training diurutkan / diranking berdasarkan masing-masing jarak paling kecil ke jarak yang paling besar kemudian dilihat untuk nilai *k*=5, *k*=7, dan *k*=9 diambil hasil konsentrasi jurusan yang lebih dominan.

Sebagai contoh perhitungan diambil data testing yang pertama yaitu data A6, berdasarkan perhitungan untuk Q1 pada data NK = A6. Pada data training dicari Ecludiance dan kemudian diurutkan berdasar atribut distance dari nilai terkecil ke terbesar seperti terlihat pada Tabel 4.

Tabel 4. Hasil Perhitungan

NK	JK	MK	KAT	RAT	REK	Distance
A69	0	6,5	70	3,2	Sistem Cerdas	90,93
A70	0	6,5	70	3,2	Jaringan	92,17
A74	0	6,5	70	3,2	Jaringan	92,68
A81	0	6,5	70	3,2	Sistem Informasi	90,97
A82	0	6,5	70	3,2	Jaringan	93,99
A86	0	6,5	70	3,2	Jaringan	91,29
A93	0	6,5	70	3,2	Sistem Cerdas	95,89
A9	1	7	71	2,98	Sistem Cerdas	101,62
A52	1	7	71	2,98	Jaringan	100,4141

Setelah diperoleh data hasil perhitungan yang telah diurutkan berdasarkan jarak terdekat kemudian di tentukan hasil nilai k, yaitu untuk k=5 (Jaringan), k= 7 (Jaringan), k=9 (Jaringan). Perhitungan ini juga dilakukan pada data testing Q1 untuk A11 s/d A40. Selanjut untuk menghitung akurasi dari masing-masing Q pada pada tiap nilai K (K3, K=5, K=7, K=9, K=11), Menghitung nilai akurasinya dapat dilakukan dengan menggunakan persamaan :

$$\text{Akurasi} = \frac{\text{Jumlah Klasifikasi benar (A2-A31)}}{\text{jumlah data uji}} \times 100\% \quad (3)$$

Hasil akurasi nilai K pada masing-masing Q1 a/d Q3 seperti terlihat pada Tabel 5.

Tabel 5. Hasil perhitungan akurasi

NO	Training (Q)	Data Testing	k=5	k=7	k=9
1	Q1	A11 s/d A40	93,40	95,23	94,60
2	Q2	A41 s/d A70	94,74	96,73	94,82
3	Q3	A71s/d A100	94,51	95,71	93,86
Nilai Akurasi Rata-rata			94,217	95,89	94,43

Berdasarkan Tabel.5 diatas nilai K yang paling besar adalah k = 7 sebesar 95,89, jadi untuk menentukan konsentrasi berdasarkan jarak maka k yang dipakai adalah K=7.

6. KESIMPULAN

Hasil *prediction* dengan Perhitungan algoritma K-NN yang diterapkan terhadap data

mahasiswa sebanyak 100 mahasiswa diperoleh hasil klasifikasi dengan prosentase akurasi sebesar 95,89 %, dengan nilai K=7 sehingga dapat disimpulkan bahwa hasil klasifikasi ini dapat digunakan untuk merekomendasikan kepada universitas PGRI dalam menentukan konsentrasi yang sesuai untuk mahasiswa.

DAFTAR PUSTAKA

- [1] Hafidh, K. (2015) Memprediksi Masa Studi Mahasiswa Menggunakan Metode Jaccard Coefficient (Studi Kasus: Mahasiswa Program Studi Teknik Informatika Jurusan Teknik Elektro Fakultas Teknik Universitas Tanjungpura).
- [2] Rohmadi, A. et al., (2015) *Cased Based Reasoning* Untuk Pemilihan Kegiatan Organisasi Mahasiswa. *Agustus Jurnal Momentum*, 17
- [3] Wahyudi, Eka. (2015) *Case-Based Reasoning* untuk Diagnosa Penyakit Jantung. Thesis. Program Studi Ilmu Komputer Fakultas MIPA Universitas Gadjah Mada.
- [4] Yulianti, W. (2016) *Aptitude Testing Berbasis Case Based Reasong* Dalam Sistem Pakar Untuk Menentukan Minat Dan Bakat Siswa Sekolah Dasar (SD).
- [5] Ulfi Saidata Aesy. (2016) Prediksi Waktu Kelulusan Mahasiswa dengan *Case Based Reasoning* pada Pascasarjana Program Studi Ilmu Komputer di Universitas Gadjah Mada.